

Non-linear optical effects and third-harmonic generation in superconductors: Cooper-pairs vs Higgs mode contribution

T. Cea,¹ C. Castellani,¹ and L. Benfatto¹

¹*ISC-CNR and Dep. of Physics, "Sapienza" University of Rome, P.le A. Moro 5, 00185, Rome, Italy**
(Dated: May 25, 2016)

The recent observation of a transmitted THz pulse oscillating at three times the frequency of the incident light paves the way to a new protocol to access resonant excitations in a superconductor. Here we show that this non-linear optical process is dominated by light-induced excitation of Cooper pairs, in analogy with a standard Raman experiment. The collective amplitude (Higgs) fluctuations of the superconducting order parameter give in general a smaller contribution, unless one designs the experiment by combining properly the light polarization with the lattice symmetry.

PACS numbers: 74.20.-z, 74.25.Gz, 74.25.N-

The enormous technological advances made in the last two decades in the time-domain spectroscopy[1, 2] pose several challenges for our understanding of the interaction of the light with the matter. The use of low-energy THz waves[3] to first excite (pump) and then measure (probe) the system is particularly interesting for superconductors, since they can access the region $\omega < 2\Delta_0$ of the optical spectrum where linear-response absorption is suppressed by the opening of a superconducting (SC) gap Δ_0 in the quasiparticle spectrum. For example, recent[4, 5] THz pump-THz probe experiments have shown that the probe field displays a periodic oscillation, whose possible connection to amplitude (Higgs) fluctuations of the SC order parameter has been investigated theoretically[7–11].

An interesting additional effect made possible by the use of intense electromagnetic (e.m.) THz field is the experimental observation[5] of the so-called third-harmonic generation (THG), i.e. the appearance below T_c in the transmitted pulse of a component oscillating three times faster than the incident light. This effect appears only below T_c with a maximum intensity at the temperature where the light frequency ω matches the SC gap value $\Delta_0(T)$, and has been attributed[5, 6] to a resonant excitation of the Higgs mode. However, we show here that THG is dominated by the resonant excitations of Cooper pairs (CP), (see Fig. 1), overlooked in previous theoretical work[5, 6].

In contrast to pump-probe experiments[4, 5], where the description of the intermediate relaxation processes of the photoexcited states becomes relevant[7–11], the THG effect can be understood as an equilibrium, non-linear optical process. In this paper we compute microscopically the non-linear optical response of a superconductor and we show that the THG essentially measures lattice-modulated *density* correlations, that in the SC state diverge at the threshold $2\Delta_0$ above which Cooper pairs (CP) proliferate. This effect induces a resonant enhancement of the THG intensity when the frequency 2ω of the incoming electric field coincides with $2\Delta_0$, as observed

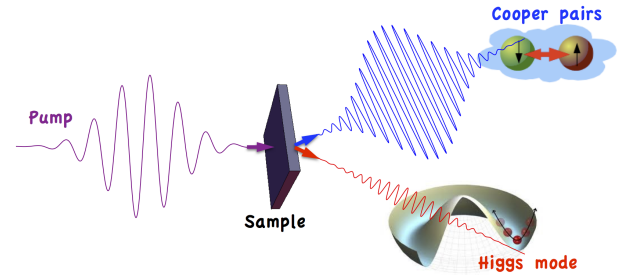


Figure 1. Schematic of the THG. An intense THz pulse shining on the SC sample generates a transmitted component oscillating three times faster, due to the resonant excitations of Cooper pairs or Higgs fluctuations. The higher intensity of the former process can be modulated by changing the polarization of the incident light.

experimentally. Once identified the relevant non-linear optical response function, we also find that the Higgs-mode contribution is largely subleading, due to symmetry reasons. Indeed, even if the Higgs mode can be excited by the THz field, as discussed previously[5, 6], it essentially decouples from the optical probe. This is a consequence of the weak coupling between the SC amplitude and density fluctuations in BCS superconductors [4, 13–15], as usually discussed in the context of Raman experiments[14, 15, 17]. The potential analogy with Raman experiments emerges also on the non-trivial dependence of the THG on the relative orientation between the e.m. field and the main crystallographic axes, due to the lattice symmetries of the band structure. This effect can be tested e.g. in cuprate superconductors, where large monocrystals have been already studied by non-linear spectroscopy[18]. Even though this polarization dependence can also be used to selectively excite the Higgs mode, its weak signal remains a major obstacle to its detection. Finally, by including the CP effects, missing in previous theoretical work[5, 6] due to an incorrect computation of the non-linear optical response, we reproduce very well the temperature dependence of the THG measured in Ref. [5].

We start from a microscopic SC model that captures the main ingredients of the problem:

$$H = \sum_{\mathbf{k}, \sigma} \xi_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} - \frac{U}{N_s} \sum_{\mathbf{q}} \Phi_\Delta^\dagger(\mathbf{q}) \Phi_\Delta(\mathbf{q}), \quad (1)$$

where $\xi_{\mathbf{k}} = \varepsilon_{\mathbf{k}} - \mu$ is the electronic dispersion with respect to the chemical potential μ , $U > 0$ is the SC coupling and $\Phi_\Delta(\mathbf{q}) = \sum_{\mathbf{k}} c_{-\mathbf{k}+\mathbf{q}/2\downarrow} c_{\mathbf{k}+\mathbf{q}/2\uparrow}$. In mean-field approximation the Green's function in the usual basis of Nambu operators $\Psi^\dagger = (c_{\mathbf{k}\uparrow}^\dagger, c_{-\mathbf{k}\downarrow})$ reads $G_0^{-1} = i\omega_n \hat{\tau}_0 - \xi_{\mathbf{k}} \hat{\tau}_3 + \Delta_0 \hat{\tau}_1$, where $\hat{\tau}_i$ are Pauli matrices, Δ_0 is the SC gap and $E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + \Delta_0^2}$. The coupling to the gauge field \mathbf{A} can be introduced by means of the Peierls substitution, $c_{i+\hat{x}}^\dagger c_i \rightarrow c_{i+\hat{x}}^\dagger c_i e^{ie\mathbf{A} \cdot \hat{x}}$. To derive the e.m. kernel we follow a standard procedure[1, 20] to derive the action S_A written in terms of the gauge field \mathbf{A} and SC collective modes. Since the coefficients of the effective action are given by fermionic susceptibilities this approach allows one to include both the quasiparticles and collective-mode contributions to the optical kernel, as it has been proven already for the linear response[3, 21]. As we shall see below, it turns out that the most relevant contributions to the non-linear current \mathbf{J}^{NL} can be written in a compact notation as

$$\mathbf{J}^{NL} \sim \mathbf{A}(\chi^{CP} + \chi^H)\mathbf{A}^2, \quad (2)$$

$$\chi^{CP} \sim \langle \rho\rho \rangle, \quad \chi^H \sim \langle \Delta\Delta \rangle. \quad (3)$$

where the CP contribution χ^{CP} probes lattice-modulated density fluctuations, while the Higgs contribution χ^H is proportional to the amplitude fluctuations. Even though both terms diverge at $\omega = \Delta_0$, the prefactor of χ^H turns out to be strongly suppressed by the particle-hole symmetry of the BCS solution, making it largely subdominant with respect to the CP one.

To make this argument quantitative we compute the non-linear response by expanding the action S_A up to the fourth order in \mathbf{A} . For an uniform field the terms relevant for the THG are then:

$$S[A] = \frac{1}{2} \sum_{\Omega_n} e^4 A_i^2(\Omega_n) \chi_{ij}^{CP}(\Omega_n) A_j^2(\Omega_n) + X_{\Delta\Delta}(\Omega_n) |\Delta(\Omega_n)|^2 + 2e^2 A_i^2(\Omega_n) \chi_{A_i^2\Delta}(\Omega_n) \Delta(-\Omega_n)$$

where $A_j^2(\Omega_n)$ is the Fourier transform of $(A_j(t))^2$ in Matsubara frequency $i\Omega_n = 2\pi nT$. The first term of Eq. (S6) is the CP response, as given by

$$\chi_{ij}^{CP}(i\Omega_n) = \langle \rho_i \rho_j \rangle = \Delta_0^2 \sum_{\mathbf{k}} \partial_i^2 \varepsilon_{\mathbf{k}} \partial_j^2 \varepsilon_{\mathbf{k}} F_{\mathbf{k}}(i\Omega_n), \quad (5)$$

$$F_{\mathbf{k}}(i\Omega_n) = \frac{1}{N_s} \frac{\tanh(E_{\mathbf{k}}/2T)}{E_{\mathbf{k}}[(i\Omega_n)^2 - 4E_{\mathbf{k}}^2]}, \quad (6)$$

where we introduced the short notation $\partial_i^2 \varepsilon_{\mathbf{k}} \equiv \partial^2 \varepsilon_{\mathbf{k}} / \partial k_i^2$. Here $\langle \dots \rangle$ denotes the correlation function for the operator $\rho_i(\mathbf{q}) = \sum_{\mathbf{k}} \partial_i^2 \varepsilon_{\mathbf{k}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}$, showing that χ_{ij}^{CP}

scales as the density-density correlation function (given in the BCS limit by $\chi_{\rho\rho} \equiv \Delta_0^2 \sum_{\mathbf{k}} F_{\mathbf{k}}(i\Omega_n)$), as anticipated in Eq. (2) above. Indeed, the band derivatives $\partial_i^2 \varepsilon_{\mathbf{k}}$ just represent in a lattice model the equivalent of the inverse mass $1/m$ for free electrons, and they always come along with a $A_i^2(\omega)$ term in the effective action[3, 20]. The second term in Eq. (S6) describes the collective fluctuations of the SC amplitude Δ , $\langle |\Delta|^2 \rangle_{\mathbf{A}=0} = 1/X_{\Delta\Delta}$, given as usual[4, 7, 13, 14, 23] by

$$X_{\Delta\Delta}(\omega) = (4\Delta_0^2 - \omega^2)F(\omega), \quad F(\omega) = \sum_{\mathbf{k}} F_{\mathbf{k}}(\omega) \quad (7)$$

By analytical continuation $i\Omega_n \rightarrow \omega + i0^+$ one can easily see that both $\chi^{CP}(\omega)$ and $F(\omega)$ display a square-root divergence as $\omega \rightarrow 2\Delta_0$, that signals the proliferation of CP above the gap. As it is well known, this effects makes the Higgs a non-relativistic mode[4], i.e. amplitude fluctuations display a overdamped resonance at $\omega = 2\Delta_0$. Finally, the third term in Eq. (S6) describes the coupling between the e.m. field and the Higgs mode, mediated by the function

$$\chi_{A_i^2\Delta}(i\Omega_n) = \langle \rho_i \Delta \rangle = 2\Delta_0 \sum_{\mathbf{k}} (\partial_i^2 \varepsilon_{\mathbf{k}}) \xi_{\mathbf{k}} F_{\mathbf{k}}(i\Omega_n). \quad (8)$$

Eq.s (S8) and (S9) define the basic correlation functions needed to compute the THG. They also explain why for BCS superconductors, where they are given explicitly by the r.h.s. of Eq. (S8)-(S9), $\chi_{A_i^2\Delta}$ is very small. Indeed, in the continuum limit, where the band dispersion can be approximated with a parabolic one $\varepsilon_{\mathbf{k}} \simeq \mathbf{k}^2/2m$, so that $\partial_i^2 \varepsilon_{\mathbf{k}} \simeq 1/m$, $\chi_{A_i^2\Delta}$ vanishes, since

$$\chi_{A_i^2\Delta}(\omega) \simeq \frac{N_F}{m} \int_{-\infty}^{\infty} d\xi \frac{\xi}{\sqrt{\xi^2 + \Delta_0^2}(\omega^2 - 4\Delta_0^2 - 4\xi^2)} \simeq 0,$$

where the integration range can be taken symmetric due to the approximate particle-hole symmetry of the BCS solution[4, 13]. This result explains the suppression of the Higgs contribution to the THG.

To derive the non-linear e.m. kernel we integrate out the amplitude fluctuations in Eq. (S6), that is equivalent to compute the RPA vertex correction of the bare bubble χ_{ij}^{CP} [6, 20]. One is then left with the action depending on the e.m. field only,

$$S^{(4)}[A] = \frac{e^4}{2} \int dt dt' \sum_{i,j} A_i^2(t) K_{ij}(t-t') A_j^2(t'), \quad (9)$$

$$K_{ij}(t-t') = [\chi_{ij}^{CP}(t-t') + \chi_{ij}^H(t-t')] \quad (10)$$

where the Higgs contribution χ_{ij}^H reads

$$\chi_{ij}^H(i\Omega_n) \equiv -\frac{\chi_{A_i^2\Delta}(i\Omega_n) \chi_{A_j^2\Delta}(i\Omega_n)}{X_{\Delta\Delta}(i\Omega_n)}, \quad (11)$$

and it is also diverging at $\omega = 2\Delta_0$ due to the vanishing of $X_{\Delta\Delta}$, see Eq. (7). The non-linear current J_i^{NL} follows

by functional derivative of Eq. (S14) with respect to \mathbf{A} :

$$J_i^{NL}(t) = -\frac{\delta S^{(4)}[A]}{\delta A_i(t)} = -2e^4 A_i(t) \int dt' \sum_j K_{ij}(t-t') A_j^2(t'), \quad (12)$$

Eq. (12), with the definition (S15) of the e.m. kernel, corresponds to Eq. (2) above. For a monochromatic incident field $\mathbf{A} = \bar{\mathbf{A}} \cos(\Omega t)$ it is given by

$$J_i^{NL}(t) = \frac{e^4 \bar{A}_i}{4} \sum_j \{ e^{-3i\Omega t} K_{ij}(2\Omega) + e^{-i\Omega t} [2K_{ij}(0) + K_{ij}(2\Omega)] + c.c. \} \bar{A}_j^2 \quad (13)$$

where one recovers the term oscillating at three times the incident frequency, with an amplitude controlled by the kernel K_{ij} evaluated at 2Ω . In the experiments of Ref. [5] the physical observable is the transmitted electric field \mathbf{E}^{tr} , that one expects to be proportional to the current (13). As a consequence the intensity of the THG can be evaluated from Eq. (13) as $I_i^{THG}(\Omega) \propto |\int dt J_i^{NL}(t) e^{3i\Omega t}|^2$, that for a monochromatic wave gives

$$I_i^{THG}(\Omega) = I_0 e^8 \bar{A}_i^2 \left| \sum_j K_{ij}(2\Omega) \bar{A}_j^2 \right|^2, \quad (14)$$

where I_0 is an overall scale factor that depends on the geometry of the experiment.

To quantify explicitly the lattice effects we compute the non-linear response for a nearest-neighbors tight-binding model on the square lattice $\varepsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y)$, and we will consider first the half-filled case $n = 1$ ($\mu = 0$), where only the SC amplitude mode contributes to the non-linear response. By making the replacement $\partial_i^2 \varepsilon_{\mathbf{k}} = 2t \cos k_i$ in Eqs. (S8) and (S9) one sees that $\chi_{A^2\Delta}$ is independent on the direction while the CP part (S8) is a tensor. Let us first consider the case of a field applied along the x axis, so that I_x^{THG} is controlled by the longitudinal K_{xx} kernel. The two separate CP ($I_x^{CP}(\Omega) \propto |\chi_{xx}^{CP}(2\Omega)|^2$) or Higgs ($I_x^H(\Omega) \propto |\chi^H(2\Omega)|^2$) contributions to the THG intensity for a monochromatic field are shown in Fig. 2a. As one can see, even if the functional form is similar for the two terms, the CP contribution is much larger, and one can roughly estimate $I_x^H \sim (\Delta/U)^4 I_{xx}^{CP}$. The predominance of the CP response implies also a non-trivial dependence of the THG intensity on the direction of the incoming applied field with respect to the crystallographic axes. In the general case of $\bar{\mathbf{A}} = A_0 \cos(\Omega t)(\cos \theta, \sin \theta)$, θ being the angle with respect to the x axis, the intensity of the transmitted pulse in the field direction is:

$$I_\theta^{THG}(\Omega) = I_0 e^8 A_0^6 |K_\theta(2\Omega)|^2 \quad (15)$$

$$K_\theta = \chi_{xx}^{CP} (\cos^4 \theta + \sin^4 \theta) + 2\chi_{xy}^C \sin^2 \theta \cos^2 \theta + \chi_{yy}^H \quad (16)$$

where we used the fact that $\chi_{yy/xy}^{CP} = \chi_{xx/xy}^{CP}$. When $\bar{\mathbf{A}}$ is applied along the diagonal ($\theta = \pi/4$) the longitudinal χ_{xx}^{CP} and transverse χ_{xy}^{CP} parts of the CP response

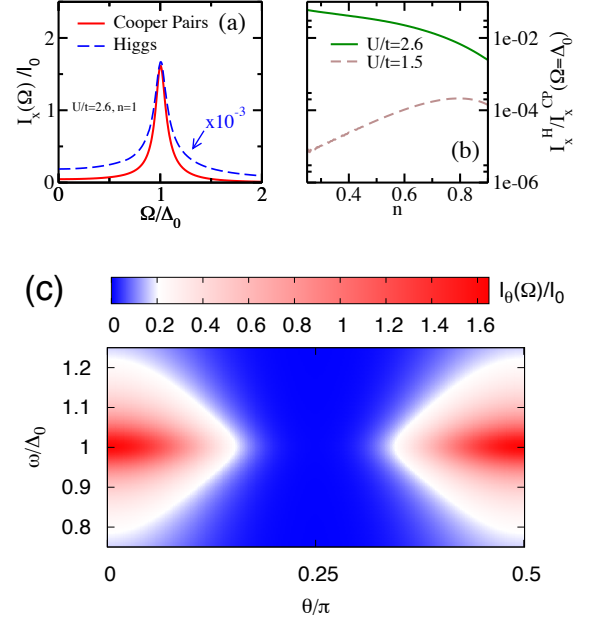


Figure 2. (a) Comparison between the CP and Higgs contribution to the THG intensity at $T = 0$ as a function of the frequency Ω for a field along x and a residual broadening $\omega + i\delta$, $\delta = 0.1\Delta_0$. (b) Relative intensity of the Higgs and CP processes at $\omega = \Delta_0$ as a function of the density for two values of the SC coupling. Here the Coulomb screening from Eq.s (S18)-(S19) has been included. (c) Map of the frequency dependence of the THG intensity, Eq. (15), for a field at arbitrary angle θ with respect to the x direction.

are equally weighted. In this peculiar configuration one sees from the definitions (S8)-(S9) that $\chi_{xx}^{CP} + \chi_{xy}^{CP} = -\Delta_0 \chi_{A^2\Delta}/2$, i.e. the diverging CP contribution cancels out, and only the resonant Higgs response remains:

$$J_{\pi/4}^{NL}(t) = \frac{e^2 \Delta_0}{U} A(t) \langle \Delta(t) \rangle, \quad (17)$$

where $\langle \Delta(t) \rangle$ is the average value of the amplitude fluctuations obtained from Eq. (S6), i.e. $\langle \Delta(\omega) \rangle = e^2 \chi_{A^2\Delta}(\omega) \mathbf{A}^2(\omega) / X_{\Delta\Delta}(\omega)$ in the frequency domain. The angular dependence of $I_\theta^{THG}(\Omega)$ for is shown in Fig. 2c as a colour map: at $\theta = \pi/4$, where one probes only the Higgs mode, the intensity is strongly suppressed, in agreement with the result shown in Fig. 2a. This prediction can be tested in systems like cuprate superconductors, where the band structure has in first approximation the symmetry discussed here and large monocrystals have been already used to probe SC resonances by pumping the system with near-infrared light[18]. It is worth noting that all these effects have been completely overlooked in the previous work[5, 6] due to the incorrect replacement in the CP term (S8) of the quantity $\partial_i^2 \varepsilon_{\mathbf{k}}$, that is *finite* at the Fermi surface, with $\xi_{\mathbf{k}}$, that is instead vanishing. This assumption removes both the divergence of the CP term $\chi_{ij}^{CP}(\omega)$ at $\omega = 2\Delta$ and its direction dependence,

and leads always to the result (S32), that is instead far from being generic. In addition, we also checked[20] that the expression (S15) can be obtained as well by means of the pseudospin formalism used in Ref. [5, 6].

Since the response function (S15) is dominated by the electronic states at the Fermi surface, the quantitative difference between the CP and Higgs response depends in general on the electron density n . To quantify this effect away from half-filling one should retain in the derivation (S6) of the effective action also the terms[20] coupling the gauge field and the Higgs mode to the phase and density fluctuations, mediated by the response functions

$$\chi_{A^2\rho}(i\Omega_n) = \langle \rho_i \rho \rangle = \Delta_0^2 \sum_{\mathbf{k}} \partial_i^2 \varepsilon_{\mathbf{k}} F_{\mathbf{k}}(i\Omega_n), \quad (18)$$

$$\chi_{\rho\Delta}(i\Omega_n) = \langle \rho \Delta \rangle = \Delta_0 \sum_{\mathbf{k}} \xi_{\mathbf{k}} F_{\mathbf{k}}(i\Omega_n). \quad (19)$$

These terms, that vanish by particle-hole symmetry in the half-filled case, are crucial to account for the screening effects of the long-range Coulomb potential, in analogy again with the known result for the Raman response function[17]. By means of straightforward but lengthy calculations one can then show[20] that the non-linear response function retains the structure (S15) with the replacements

$$\chi_{ij}^{CP} \rightarrow \chi_{ij}^{CP,sc} = \chi_{ij}^{CP} - \frac{\chi_{A^2\rho}^2}{\chi_{\rho\rho}}, \quad (20)$$

$$\chi^H \rightarrow \chi^{H,sc} = -\frac{(\chi_{A^2\Delta} - \chi_{A^2\rho}\chi_{\rho\Delta}/\chi_{\rho\rho})^2}{X_{\Delta\Delta} - \chi_{\rho\Delta}^2/\chi_{\rho\rho}}. \quad (21)$$

where we used the fact that for the lattice model under consideration the function (18) is isotropic in the spatial indexes. While the mixing to the density and phase modes does not affect[4] the pole of the Higgs, identified now by the vanishing of the denominator of Eq. (S19), it is crucial to screen both the CP and Higgs response as one moves away from half-filling. In Fig. 2b we show the ratio $I_x^{CP}(\Delta)/I_x^H(\Delta)$ as a function of the electron density n for two values of the SC coupling U . As we can see, even for the large value $U/t = 2.6$ of the SC coupling, where the deviations from the BCS (approximate) particle-hole symmetric case become more prominent, inducing a larger coupling of the Higgs to the light, the CP part remains the predominant one for the longitudinal response even in the low-density regime.

In addition to the strong direction dependence of the THG intensity, a second check of the origin of the THG effect is its temperature evolution, measured in Ref. [5]. For a polycrystalline sample one should average the kernel Eq. (16) over θ , to account for the random direction of the e.m. field with respect to the crystallographic axes. One then finds that $\bar{J}_\theta^{NL} = [J_x^{NL} + J_{\pi/4}^{NL}]/2 \simeq J_x^{NL}/2$, so that one expects that the CP processes dominate. To check this we compute[20] the non-linear current induced

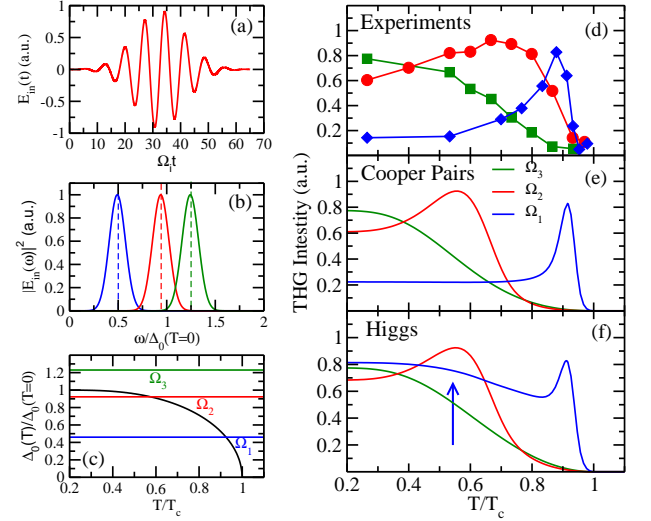


Figure 3. (a) Profile of the incoming field, and (b) corresponding power spectra, used to simulate the experiments of Ref. [5]. The three central frequencies Ω_i of the power spectra are compared in panel (c) to the temperature dependence of the SC gap. For each Ω_i the resonant condition $\Omega_i = \Delta_0(T)$ occurs at a different temperature, in analogy with Ref. [5]. (d-e) Temperature evolution of the THG intensity in the experiments (d), in the case of CP processes (e) and in the case of the Higgs processes alone (f), computed for $n = 1, U/t = 2.6$. Data computed at different Ω_i are normalized to have a similar overall scale, as done in panel (d). The excitation of the Higgs mode alone gives the wrong T dependence of the THG signal for the lowest frequency Ω_1 , as marked by the arrow.

by an incoming electric field $\mathbf{A}(t)$ having a wave-packet profile similar to the one used in the experiments of Ref. [5], see Fig. 3a. In the frequency domain this wave packet corresponds to the power spectra shown in Fig. 3b, centred at three possible values Ω_i of the incoming frequency. The temperature evolution of the THG intensity, i.e. $I_x(3\Omega_i)$, is then shown in Fig. 3d-f, where we compare the experimental data from Ref. [5] (panel d) with the theoretical calculations done including only the CP processes (panel e) or the Higgs contribution only (panel f). Apart from the small overall intensity of the THG Higgs signal, that cannot be seen in the normalized data of Fig. 3, the excitation of the Higgs mode alone fails to reproduce the temperature dependence of the signal at the lowest frequency Ω_1 .

In conclusion, we studied the non-linear optical effects responsible for the THG in a superconductor. Since the relevant response function (S15) measures lattice-modulated density fluctuations, see Eq. (S8), the optical process responsible for the THG is equivalent to a resonant excitation of CP. The Higgs-mode contribution is instead much smaller, since its coupling (S9) to the optical probe is suppressed by symmetry, in analogy with standard Raman experiment[13–15, 17], unless some addi-

tional channel makes the Higgs Raman visible[13, 14, 24]. In addition, also for THG experiments one can orient the light polarization with respect to the main crystallographic axes in order to modulate the THG intensity due to Cooper pairs. Even though this effect can be used in principle to selectively excite the Higgs signal, the weakness of its coupling to the light represents a major obstacle to its detection in optical experiments, in analogy with the results recently discussed in the context of linear optical spectroscopy[4]. The interplay between the non-linear optical effects discussed here and the non-equilibrium processes[7–11] addressed in the context of pump-probe experimental protocols[1, 2, 4, 5] remains an open question, that certainly deserves future experimental and theoretical work.

This work has been supported by Italian MIUR under projects FIRB-HybridNanoDev-RBFR1236VV, PRINRIDEIRON-2012X3YFZ2 and Premiali-2012 AB-NANOTECH.

* lara.benfatto@roma1.infn.it

- [1] J. Orenstein, Phys. Today **65**, 44 (2012).
- [2] C. Giannetti, M. Capone, D. Fausti, M. Fabrizio and F. Parmigiani, *Ultrafast optical spectroscopy of strongly correlated materials and high-temperature superconductors: a non-equilibrium approach*, preprint 2015.
- [3] For a review, see, e.g., T. Kampfrath, K. Tanaka, K. A. Nelson, Nat. Photonics **7**, 680, (2013).
- [4] Ryusuke Matsunga and Ryo Shimano, Phys. Rev. Lett.**109**, 187002 (2012).
- [5] Ryusuke Matsunaga, Yuki I. Hamada, Kazumasa Makise, Yoshinori Uzawa, Hirotaka Terai, Zhen Wang, and Ryo Shimano, Phys. Rev. Lett.**111**, 057002 (2013).
- [6] Ryusuke Matsunaga, Naoto Tsuji, Hiroyuki Fujita, Arata Sugioka, Kazumasa Makise, Yoshinori Uzawa, Hirotaka Terai, Zhen Wang, Hideo Aoki, Ryo Shimano, Science **345**, 1145 (2014).
- [7] A. F. Volkov and S. M. Kogan, Zh. Eksp. Teor. Fiz. 65, 2038 (1973) [Sov. Phys. JETP 38, 1018 (1974)].
- [8] R. A. Barankov, L. S. Levitov, and B. Z. Spivak, Phys. Rev. Lett. **93**, 160401 (2004).
- [9] R. A. Barankov and L. S. Levitov, Phys. Rev. Lett. **96**, 230403 (2006).
- [10] H. Krull, D. Manske, G.S. Uhrig, A.P. Schnyder, Phys. Rev. B**90**, 014515 (2014).
- [11] T. Papenkort, T. Kuhn, and V. M. Axt, Phys. Rev. B **78**, 132505 (2008).
- [12] N. Tsuji and H. Aoki, PRB **92**, 064508 (2015).
- [13] P. Littlewood and C. M. Varma, Phys. Rev. B**26**, 4883 (1982).
- [14] T. Cea and L. Benfatto, Phys. Rev. B **90**, 224515 (2014).
- [15] D. Pekker and C. M. Varma, Annual Reviews of Condensed Matter Physics Volume 6, 269 (2015).
- [16] T. Cea, C. Castellani, G. Seibold, L. Benfatto, Phys. Rev. Lett.**115**, 157002 (2015).
- [17] T. P. Devereaux and R. Hackl, Rev. Mod. Phys. **79**, 175 (2007).
- [18] B. Mansart, J. Lorenzana, A. Manna, A. Odeh, M. Scarongella, M. Cherguib, and F. Carbone, Proc. Nat. Acad. Sc. **110**, 4539 (2013).
- [19] N. Nagaosa, *Quantum Field Theory in Condensed Matter Physics*, Springer, 1999.
- [20] For more details see Supplementary Material at..
- [21] A. Paramekanti, M. Randeria, T.V. Ramakrishnan, and S.S. Mandhal, Phys. Rev. B **62**, 6786 (2000).
- [22] See e.g. L. Benfatto, A. Toschi, and S. Caprara, Phys. Rev. B. **69**, 184510 (2004) and references therein.
- [23] I. O. Kulik et al., J. Low. Temp. Phys. **43**, 591 (1981).
- [24] M.-A. Méasson, Y. Gallais, M. Cazayous, B. Clair, P. Rodière, and L. Cario, and A. Sacuto Phys. Rev. B **89**, 060503(R) (2014).

Supplemental Material

DERIVATION OF THE NON-LINEAR OPTICAL KERNEL

Let us start from Eq. (3) of the manuscript, that can be rewritten in real space as:

$$H = -t \sum_{\langle i,j \rangle \sigma} (c_{i\sigma}^\dagger c_{j\sigma} + h.c.) - U \sum_i c_{i\uparrow}^\dagger c_{i\downarrow}^\dagger c_{i\downarrow} c_{i\uparrow}. \quad (S1)$$

To investigate the physics of the collective fluctuations around the mean field solution we follow the usual Hubbard-Stratonovich (HS) procedure[S1], as implemented e.g. in Refs. S2 and S3. We then introduce in the action for the fermions a bosonic complex field $\psi_\Delta(\tau)$ which decouples the onsite interaction term of (S1) in the pairing channel. At $T < T_c$ one can choose to represent the superconducting (SC) fluctuations in polar (amplitude and phase) coordinates, by decomposing $\psi_\Delta(\tau) = [\Delta_0 + \Delta_i(\tau)]e^{i\theta_i(\tau)}$, where $\Delta_i(\tau)$ represent the amplitude fluctuations of ψ_Δ around the mean-field value Δ_0 of the SC order parameter and θ its phase fluctuations. By making a Gauge transformation $c_i \rightarrow c_i e^{i\theta_i/2}$ the dependence on the phase degrees of freedom is made explicit in the action. We also add to the model (S1) an additional interaction term representing the effect of Coulomb interactions:

$$H_c = \frac{1}{2} \sum_{\substack{\mathbf{k}, \mathbf{k}', \mathbf{q} \\ \sigma, \sigma'}} V(\mathbf{q}) c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} c_{\mathbf{k}, \sigma} \quad (S2)$$

where $V(\mathbf{q})$ is the Fourier transform of the Coulomb potential in the D dimensional lattice. At small \mathbf{q} it re-

duces to the expression in the continuum limit, so that $V(\mathbf{q}) \rightarrow \lambda e^2/|\mathbf{q}|^{D-1}$ where $\lambda = 4\pi/\epsilon_B$ for $D = 3$ while $\lambda = 2\pi/\epsilon_B$ for $D = 2$, ϵ_B being the background dielectric constant. This term can be decoupled in the particle-hole channel by means of an additional HS field $\psi_\rho = \rho_o + \rho$, which couples to the electronic density $\Phi_{\rho,i} = \sum_\sigma c_{i\sigma}^\dagger c_{i\sigma}$ and represents the density fluctuations ρ of the system around the mean-field value ρ_0 . Finally, the Gauge field \mathbf{A} can be introduced by means of the Peierls substitution $c_{i+\hat{x}}^\dagger c_i \rightarrow c_{i+\hat{x}}^\dagger c_i e^{ie\mathbf{A} \cdot \hat{x}}$, that modifies only the kinetic part of the Hamiltonian.

After the Hubbard-Stratonovich decoupling the action is quadratic in the fermionic fields that can then be integrated out leading to the effective action for the fields Δ , θ , ρ and \mathbf{A} :

$$S_{eff}[\Delta, \theta, \rho] = S_{MF} + S_{FL}[\Delta, \theta, \rho, \mathbf{A}] \quad , \quad (S3)$$

where $S_{MF} = \frac{N\Delta_0^2}{TU} + \frac{N\rho_0^2}{TU} - \text{Tr} \ln(-G_0^{-1})$ is the mean-field action, $G_0^{-1} = i\omega_n \hat{\sigma}_0 - \xi_{\mathbf{k}} \hat{\sigma}_3 + \Delta_0 \hat{\sigma}_1$ is the BCS Green's function and

$$S_{FL} = \sum_{n \geq 1} \frac{\text{Tr}(G_0 \Sigma)^n}{n} \quad (S4)$$

is the fluctuating action, with the trace acting both in spin and momentum space. Here $\Sigma_{kk'}$ denotes the self-energy for the fluctuating fields, which reads explicitly:

$$\begin{aligned} \Sigma_{kk'} = & -\sqrt{\frac{T}{N}} \Delta(k-k') \sigma_1 - \sqrt{\frac{T}{N}} \rho(k-k') \sigma_3 - \sqrt{\frac{T}{N}} \frac{i}{2} \theta(k-k') [(k-k')_0 \sigma_3 - (\xi_{\mathbf{k}} - \xi_{\mathbf{k}'}) \sigma_0] - \\ & - \frac{T}{2N} \sum_{q_i, i} \theta(q_1) \theta(q_2) \frac{\partial^2 \xi_{\mathbf{k}}}{\partial k_i^2} \sin \frac{\mathbf{q}_{1,i}}{2} \sin \frac{\mathbf{q}_{2,i}}{2} \sigma_3 \delta(q_1 + q_2 - k + k') + \\ & + \sum_i \left[A_i(\omega - \omega') \frac{\partial \xi_{\mathbf{k}}}{\partial k_i} \sigma_0 + \frac{1}{2} A_i^2(\omega) \frac{\partial^2 \xi_{\mathbf{k}}}{\partial k_i^2} \sigma_3 + \frac{1}{3!} A_i^3(\omega) \frac{\partial^3 \xi_{\mathbf{k}}}{\partial k_i^3} \sigma_0 + \frac{1}{4!} A_i^4(\omega) \frac{\partial^4 \xi_{\mathbf{k}}}{\partial k_i^4} \sigma_3 \right] \end{aligned} \quad (S5)$$

with $k = (i\Omega_n, \mathbf{k})$ and $\Omega_n = 2\pi T n$ bosonic Matsubara frequencies. Here $A_j^2(\omega) \equiv \int d\omega' A_j(\omega - \omega') A_j(\omega')$ is the Fourier transform of $[A(t)]^2$, and analogous convolution formulae hold for higher powers of the gauge field. The second line of Eq. (S5) represents the transcription on the lattice of the usual $(\nabla \theta)^2$ term for a continuum

model, analogously to the $A_i^2(\omega)$ term that represents the transcription of the usual diamagnetic term $\mathbf{A}^2 n/m$ in the continuum. In addition, in contrast to the continuum model, the lattice self-energy (S5) depends in principle[S2, S3] on all higher-order powers of the θ and \mathbf{A} fields. Here however we only retained the terms relevant for the derivation of the action up to terms of order

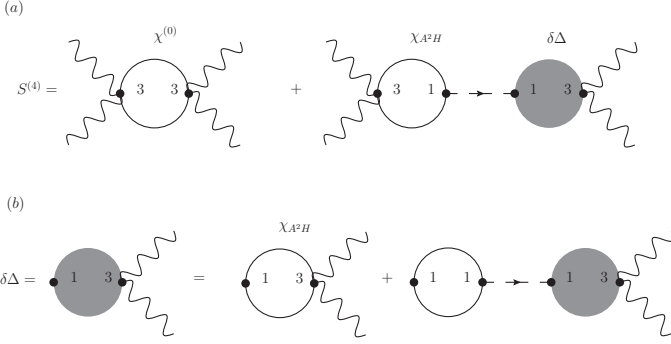


Figure S1. Figure (a): diagrammatic representation of $S^{(4)}$. The wavy and dashed lines denote the gauge and Higgs fields, respectively, while the labels 1, 3 refer to the vertex insertions of the Pauli matrices σ_1 and σ_3 , respectively. Figure (b) represents the vertex correction of $\chi_{A^2\Delta}$ in the amplitude channel, which defines the variation of the order parameter, $\delta\Delta$, due to the applied external field \mathbf{A} .

A_i^4 , and we considered directly the case of an uniform e.m. field.

As one can see from Eq. (S4) the effective action gives an expansion on powers of the bosonic fields, whose coefficients are fermionic susceptibilities that contain all the

relevant information on the quasiparticle degrees of freedom. Away from half-filling the phase/density fluctuations are in general coupled both to the gauge field and to the amplitude fluctuations Δ . The expansion of S up to quartic order in \mathbf{A} is then given by the generalization of Eq. (5) of the manuscript, i.e.

$$S[A] = \frac{1}{2} \sum_{\Omega_n} e^4 A_i^2(\Omega_n) \chi_{ij}^{CP}(\Omega_n) A_j^2(\Omega_n) + 2e^2 A_i^2(\Omega_n) \chi_{A_i^2\Delta}(\Omega_n) \Delta(-\Omega_n) + 2e^2 A_i^2(\Omega_n) \chi_{A_i^2\rho}(\Omega_n) [\rho(-\Omega_n) - i\Omega_n \theta(-\Omega_n)] + \frac{1}{2} \sum_q \Psi^\dagger(q) \hat{M}_{FL}(q) \Psi(q), \quad (\text{S6})$$

where $\Psi^T(q) = (\Delta(q) \ \theta(q) \ \rho(q))$ is a vector containing the fluctuating fields, whose Gaussian fluctuations are described by the matrix[S4]:

$$\hat{M}_{FL} = \begin{pmatrix} 2/U + \chi_{\Delta\Delta}(\Omega_n) & \frac{i\Omega_n}{2} \chi_{\rho\Delta}(\Omega_n) & \chi_{\rho\Delta}(\Omega_n) \\ -\frac{i\Omega_n}{2} \chi_{\rho\Delta}(-\Omega_n) & \frac{\Omega_n^2}{4} \chi_{\rho\rho}(\Omega_n) & -\frac{i\Omega_n}{2} \chi_{\rho\rho}(\Omega_n) \\ \chi_{\rho\Delta}(-\Omega_n) & \frac{i\Omega_n}{2} \chi_{\rho\rho}(\Omega_n) & -1/V_{\mathbf{q}} + \chi_{\rho\rho}(\Omega_n) \end{pmatrix}. \quad (\text{S7})$$

The various bubbles, given explicitly in the main manuscript, are defined in terms of the Green's functions, as

$$\chi_{ij}^{CP}(i\Omega_n) \equiv \frac{T}{N_s} \sum_{\mathbf{k}, m} \partial_i^2 \varepsilon_{\mathbf{k}} \partial_j^2 \varepsilon_{\mathbf{k}} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_3 G_0(\mathbf{k}, i\omega_m) \sigma_3] \quad (\text{S8})$$

$$\chi_{A_i^2\Delta}(i\Omega_n) \equiv \frac{T}{N_s} \sum_{\mathbf{k}, m} \partial_i^2 \varepsilon_{\mathbf{k}} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_3 G_0(\mathbf{k}, i\omega_m) \sigma_1] \quad (\text{S9})$$

$$\chi_{A_i^2\rho}(i\Omega_n) \equiv \frac{T}{N_s} \sum_{\mathbf{k}, m} \partial_i^2 \varepsilon_{\mathbf{k}} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_3 G_0(\mathbf{k}, i\omega_m) \sigma_3] \quad (\text{S10})$$

$$\chi_{\rho\Delta}(i\Omega_n) \equiv \frac{T}{N_s} \sum_{\mathbf{k}, m} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_3 G_0(\mathbf{k}, i\omega_m) \sigma_1] \quad (\text{S11})$$

$$\chi_{\rho\rho}(i\Omega_n) \equiv \frac{T}{N_s} \sum_{\mathbf{k}, m} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_3 G_0(\mathbf{k}, i\omega_m) \sigma_3] \quad (\text{S12})$$

$$X_{\Delta\Delta}(i\Omega_n) \equiv \frac{2}{U} + \chi_{\Delta\Delta} = \frac{2}{U} + \frac{T}{N_s} \sum_{\mathbf{k}, m} \text{Tr} [G_0(\mathbf{k}, i\omega_m + i\Omega_n) \sigma_1 G_0(\mathbf{k}, i\omega_m) \sigma_1] \quad (\text{S13})$$

As one can see, apart from the modulation factors $\partial_i^2 \varepsilon_{\mathbf{k}}$,

one recovers that $\chi_{A_i^2\Delta} \sim \chi_{\rho\Delta}$ and $\chi_{A_i^2\rho} \sim \chi_{\rho\rho}$. While in

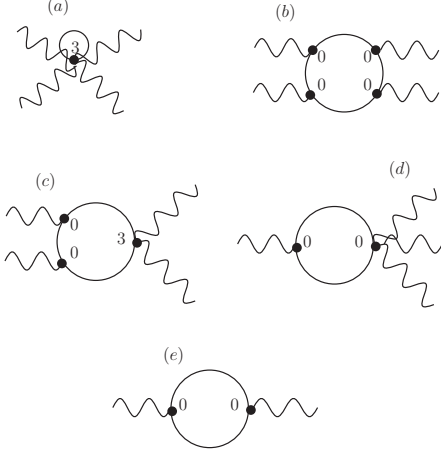


Figure S2. (a)-(d) Additional diagrams contributing to $S^{(4)}$. The (a) is independent on the transferred frequencies, while the remaining ones (b)-(d) are trivially zero when computed at zero external momenta. (e) Diagram defining the paramagnetic response in the linear response regime. It leads to the emptying of the condensate in the static limit ($\omega = 0, \mathbf{q} \rightarrow 0$), whereas it vanishes at $\mathbf{q} = 0$ and finite ω .

the low-density limits the derivatives of the band dispersions just reduces to a constant, at half-filling ($\mu = 0$), where particle-hole symmetry is exact, they are crucial to decouple the amplitude mode and the e.m. field from the density/phase modes. In this case one only retains the second line of Eq. (S6), and after integration of the Higgs mode one is left with Eq. (10) of the manuscript,

$$S^{(4)}[A] = \frac{e^4}{2} \int dt dt' \sum_{i,j} A_i^2(t) K_{ij}(t-t') A_j^2(t') \quad (\text{S14})$$

$$K_{ij}(t-t') = [\chi_{ij}^{CP}(t-t') + \chi_{ij}^H(t-t')] \quad (\text{S15})$$

where

$$\chi_{ij}^H(i\Omega_n) \equiv -\frac{\chi_{A_i^2\Delta}(i\Omega_n)\chi_{A_j^2\Delta}(i\Omega_n)}{X_{\Delta\Delta}(i\Omega_n)}, \quad (\text{S16})$$

The structure of the quartic action (S14) is diagrammatically represented in fig. S1. Here $\delta\Delta$ denotes the variation of the order parameter from its equilibrium value due to the external perturbation, which is obtained by dressing χ_{A^2H} with the vertex correction in the amplitude channel (see fig. S1-(b)):

$$\delta\Delta(\omega) = e^2 \frac{\chi_{A^2\Delta}(\omega)}{X_{\Delta\Delta}(\omega)} \mathbf{A}^2(\omega). \quad (\text{S17})$$

Notice that in addition to the diagrams shown in Fig. S1, and included in the Eq. (S14), one can have in principle several other terms of order A_i^4 , coming from the insertion of various A_i^n term of the self-energy (S5), as shown in Fig. S2 (a)-(d). They have been omitted in $S^{(4)}$ since the first one is independent on the transmitted

frequencies, while those having the σ_0 insertions trivially vanish when computed at zero external momenta. This is indeed a general result which follows from elementary algebra principles and holds for the whole class of diagrams having an arbitrary number of insertions of σ_0 and only one insertion of σ_i (with $i = 0, \dots, 3$). Despite such a rule does not hold any more in the presence of impurities, we expect that all this kind of diagrams will still be regular functions of the external frequencies. An illustrative example is represented by the diagram S2-(e), which defines the paramagnetic response in the linear response regime. As it is well known[S3], in the static limit ($\omega = 0, \mathbf{q} \rightarrow 0$) it accounts for the reduction of the superfluid stiffness at finite temperature due to thermal excitations of quasiparticles, while it vanishes in the opposite dynamic limit ($\mathbf{q} = 0, \omega \neq 0$). In the presence of disorder it also contributes to the dynamic limit, leading to the optical absorption at $\omega > 2\Delta_0$, without however any divergence at $\omega = 2\Delta_0$, in contrast to the diagrams constructed with two σ_3 insertions.

Away from half-filling the phase/density fluctuations couple to the fields Δ and A . In particular, the density fluctuations ρ , mediating the Coulomb repulsion, screen the other fields. By expliciting integrating the density and phase fields from Eq. (S6), and taking the limit $\mathbf{q} \rightarrow 0$ where $1/V_{\mathbf{q}} \rightarrow 0$, one is left with the same structure (S15) of the non-linear kernel, provided that the CP and Higgs part as screened as given by Eqs. (19)-(20) of the main text, i.e.

$$\chi_{ij}^{CP} \rightarrow \chi_{ij}^{CP,sc} = \chi_{ij}^{CP} - \frac{\chi_{A^2\rho}^2}{\chi_{\rho\rho}}, \quad (\text{S18})$$

$$\chi^H \rightarrow \chi^{H,sc} = -\frac{(\chi_{A^2\Delta} - \chi_{A^2\rho}\chi_{\rho\Delta}/\chi_{\rho\rho})^2}{X_{\Delta\Delta} - \chi_{\rho\Delta}^2/\chi_{\rho\rho}}. \quad (\text{S19})$$

THG BY AN IMPULSIVE ELECTRIC FIELD

In the previous section we have studied the THG when the applied electric field is a monochromatic wave. To better simulate the experiment described in [S5] we now consider an impulsive multi-cyclic field, having a relatively narrow spectrum of frequencies. To this purpose we choose:

$$A(t) = A_0 e^{-(t/\tau)^2} \cos(\Omega t), \quad (\text{S20})$$

where τ is a time constant which determines the duration of the impulse. The incoming electric field, given by the time derivative of Eq. (S20), is shown in Fig. 3a of the manuscript, and its corresponding power spectrum is shown in Fig. 3b for three different values of the central frequency $\Omega = \Omega_i$. By using the definition of the non-linear current given in Eq. (14) of the main text one easily finds that:

$$J_{\theta}^{NL}(\omega) = -\frac{e^4 A_0^3 \tau^2}{\sqrt{8}} \int d\omega' K_{\theta}(\omega') \exp \left[-\frac{\tau^2}{8} (3\omega'^2 + 2\omega^2 - 4\omega\omega' + 2\Omega^2) \right] \cosh \left[\frac{\tau^2}{2} \Omega (\omega - \omega') \right] \left[1 + e^{-\frac{\tau^2}{2} \Omega^2} \cosh \left(\frac{\tau^2}{2} \Omega \omega' \right) \right]. \quad (S21)$$

As we discuss in the main text, the outgoing electric field is expected to be proportional to the non-linear current, $|E_{out}(\omega)|^2 \propto |J_x^{NL}(\omega)|^2$. The power spectra of the outgoing fields generated by the CP or Higgs processes are shown in Fig. S3 for the lowest (Ω_1) and intermediate (Ω_2) value of the central frequency. As one can see, in addition to the peak at $\omega = \Omega_i$ an additional peak appears around $\omega = 3\Omega_i$, whose absolute intensity and temperature variation depend on the nature of the resonant process. Notice the close resemblance between the experimental data shown in Fig. 3D of Ref. [S5] and the temperature evolution of the power spectra for $\Omega = \Omega_2$, where the variation of THG intensity is accompanied by a small shift of the peak maximum. To make a direct comparison with the data shown in Ref. [S5], we adopt the same definition of THG intensity at each temperature as the one given there, i.e. we take the value at $\omega = 3\Omega$ ($I_x^{THG}(\Omega) = I_0 |J_x^{NL}(3\Omega)|^2$) and we plot it as a function of temperature in Fig. 3e-f of the manuscript. As it is clear from Fig. S3a-b, while for $\Omega = \Omega_2$ the temperature evolution of the THG intensity is similar for CP and Higgs processes, for $\Omega = \Omega_1$ the relative enhancement of the CP intensity at the temperature $T \simeq 0.9T_c$, where the resonant condition $\Omega_2 = \Delta_0(T)$ occurs, is much more pronounced with respect to the Higgs processes. When all the intensities are rescaled to the value obtained at the resonance, as done in Fig. 3d-f of the main text, this difference leads to the different T evolution of the $I^{THG}(T)$ shown in the two panels e and f.

DERIVATION OF THE NON-LINEAR OPTICAL KERNEL WITHIN THE PSEUDOSPIN FORMALISM

Let us shown how the non-linear optical kernel K_{ij} of Eq. (S15) can be derived also by means of the Anderson pseudospin formalism, used in the previous theoretical work[S5, S6]. The Anderson pseudospin are defined by $\sigma_{\mathbf{k}} = \frac{1}{2} \Psi^\dagger \sigma \Psi_{\mathbf{k}}$, where $\Psi_{\mathbf{k}} = (c_{\mathbf{k}\uparrow}, c_{-\mathbf{k}\downarrow}^\dagger)^T$ is the Nambu spinor and $\sigma = (\sigma_1, \sigma_2, \sigma_3)$, so that the components of $\sigma_{\mathbf{k}}$ represent the amplitude, phase and density degrees of freedom, respectively. The mean-field Hamiltonian in the presence of the (uniform) electromagnetic field can then be written as:

$$H_{BCS} = 2 \sum_{\mathbf{k}} \mathbf{b}_{\mathbf{k}} \cdot \sigma_{\mathbf{k}} \quad , \quad (S22)$$

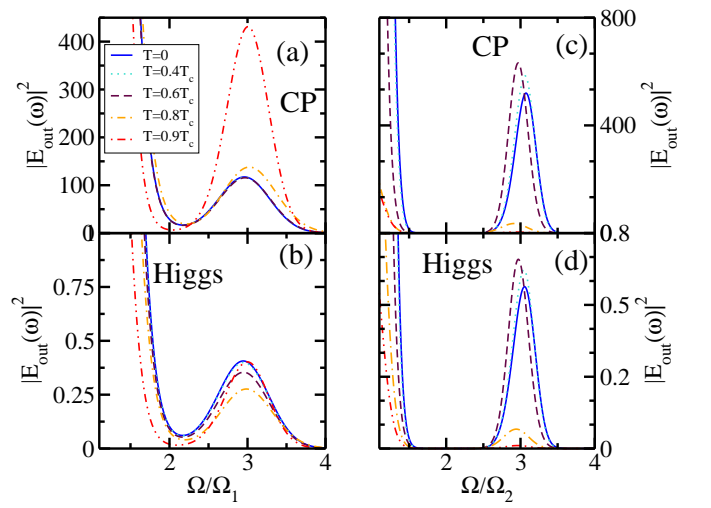


Figure S3. Temperature evolution of the power spectra of the outgoing electric field generated by CP (a,c) or Higgs (b,d) processes corresponding to the incoming THz pulse (S20) for two different values of the central frequency Ω . The peak around $3\Omega_i$ defines the THG intensity. The calculations have been done in the case $n = 1, U/t = 2.6$. Simulations at different electron densities give similar results for the relative T variations.

with:

$$\mathbf{b}_{\mathbf{k}} = \left(-\Delta'(t), -\Delta''(t), \frac{\xi_{\mathbf{k}-\mathbf{e}\mathbf{A}(t)} + \xi_{\mathbf{k}+\mathbf{e}\mathbf{A}(t)}}{2} \right) \quad , \quad (S23)$$

and $\Delta(t) = \frac{U}{N_s} \sum_{\mathbf{k}} \langle \sigma_{\mathbf{k}}^x(t) + i\sigma_{\mathbf{k}}^y(t) \rangle$ the SC order parameter. Here the brackets $\langle \cdot \rangle$ denote the expectation value computed over the BCS ground state $|\Psi_{BCS}\rangle$, so that the previous definition of Δ is self-consistent.

The equation of motion for $\langle \sigma_{\mathbf{k}}(t) \rangle$, as deduced from the Shroedinger equation for $|\Psi_{BCS}(t)\rangle$, can be written as:

$$\partial_t \langle \sigma_{\mathbf{k}} \rangle = 2\mathbf{b}_{\mathbf{k}} \times \langle \sigma_{\mathbf{k}} \rangle \quad . \quad (S24)$$

We solve (S24) in perturbation theory with respect to A^2 , by writing:

$$\langle \sigma_{\mathbf{k}}(t) \rangle = \langle \sigma_{\mathbf{k}}(0) \rangle + \delta \sigma_{\mathbf{k}}(t) \quad \text{and} \quad \mathbf{b}_{\mathbf{k}}(t) = \mathbf{b}_{\mathbf{k}}(0) + \delta \mathbf{b}_{\mathbf{k}}(t) \quad , \quad (S25)$$

with:

$$\langle \sigma_{\mathbf{k}}(0) \rangle = \left(\frac{\Delta_0}{2E_{\mathbf{k}}}, 0, -\frac{\xi_{\mathbf{k}}}{2E_{\mathbf{k}}} \right) \quad ; \quad (S26a)$$

$$\mathbf{b}_{\mathbf{k}}(0) = (-\Delta_0, 0, \xi_{\mathbf{k}}) \quad ; \quad (S26b)$$

$$\delta \mathbf{b}_{\mathbf{k}}(t) = \left(-\delta \Delta'(t), 0, \frac{e^2}{2} \sum_i \partial_i^2 \epsilon_{\mathbf{k}} A_i^2(t) + O(A^4) \right), \quad (\text{S26c})$$

where we included for the sake of simplicity only the amplitude fluctuations $\delta \Delta'(t)$, the extension to the case having also phase fluctuations $\delta \Delta''(t)$ being straightforward. The linearized equations of motion at $T = 0$ then read:

$$\begin{cases} \partial_t \delta \sigma_{\mathbf{k}}^x(t) = -2\xi_{\mathbf{k}} \delta \sigma_{\mathbf{k}}^y(t) \\ \partial_t \delta \sigma_{\mathbf{k}}^y(t) = 2\xi_{\mathbf{k}} \delta \sigma_{\mathbf{k}}^x(t) + 2\Delta_0 \delta \sigma_{\mathbf{k}}^x(t) + \\ + \frac{1}{2E_{\mathbf{k}}} [e^2 \Delta_0 \sum_i \partial_i^2 \epsilon_{\mathbf{k}} A_i^2(t) - 2\xi_{\mathbf{k}} \delta \Delta'(t)] \\ \partial_t \delta \sigma_{\mathbf{k}}^z(t) = -2\Delta_0 \delta \sigma_{\mathbf{k}}^y(t) \end{cases}. \quad (\text{S27})$$

Notice that (S27) have to be solved self-consistently, by imposing: $\delta \Delta(t) = \frac{U}{N_s} \sum_{\mathbf{k}} [\delta \sigma_{\mathbf{k}}^x(t) + i \delta \sigma_{\mathbf{k}}^y(t)]$.

The first and the last equations of the system (S27), along with the initial conditions $\delta \sigma_{\mathbf{k}}(0) = 0$, lead to the identity:

$$\xi_{\mathbf{k}} \delta \sigma_{\mathbf{k}}^z(t) \equiv \Delta_0 \delta \sigma_{\mathbf{k}}^x(t), \quad (\text{S28})$$

which allows one to reduce the number of equations.

The non-linear current can be expressed in the pseudospin formalism as[S5]:

$$J_i^{NL} = -2e^2 A_i(t) \sum_{\mathbf{k}} \frac{\partial^2 \epsilon_{\mathbf{k}}}{\partial k_i^2} \delta \sigma_{\mathbf{k}}^z(t), \quad (\text{S29})$$

i.e. it is controlled by density fluctuations, in agreement with the general argument discussed in the main manuscript. To compute the current one then needs to deduce $\delta \sigma_{\mathbf{k}}^z(t)$ by solving the system (S27). By simple algebra one can express it in the Fourier space as:

$$\begin{aligned} \delta \sigma_{\mathbf{k}}^z(\omega) = e^2 \sum_j \partial_j^2 \epsilon_{\mathbf{k}} \frac{\Delta_0^2}{E_{\mathbf{k}}(\omega^2 - 4E_{\mathbf{k}}^2)} A_j^2(\omega) + \\ - \frac{\Delta_0 \xi_{\mathbf{k}}}{E_{\mathbf{k}}(\omega^2 - 4E_{\mathbf{k}}^2)} \delta \Delta'(\omega), \end{aligned} \quad (\text{S30})$$

where $\delta \Delta'(\omega)$ from Eq. (S27) coincides with the expression (S17) derived above from Eq. (S6). By inserting

then Eq. (S17) into Eq. (S29) one finds back the definition (S15) of the electromagnetic kernel. In particular, one recovers again the presence of a CP term that is strongly diverging at $\omega = 2\Delta_0$. This contribution has been overlooked by the authors of Refs. [S5, S6] since they made the replacement

$$\sum_{\mathbf{k}} \frac{\partial^2 \epsilon_{\mathbf{k}}}{\partial k_i^2} \delta \sigma_{\mathbf{k}}^z(t) \rightarrow -\frac{1}{2} \sum_{\mathbf{k}} \xi_{\mathbf{k}} \delta \sigma_{\mathbf{k}}^z(t). \quad (\text{S31})$$

By making the assumption (S31) and using the identity (S28), the authors of Ref. [S5, S6] deduced that the non-linear current should be always written as:

$$J^{NL}(t) = \frac{e^2 \Delta_0}{U} A(t) \langle \Delta(t) \rangle. \quad (\text{S32})$$

However, the replacement (S31) is clearly *wrong*, since from Eq.s (S29) and (S30) one immediately sees that J_i^{NL} , being a *non-linear* current, depends in general in asymmetric way on the two A_i components. As we discuss in the main text, within the formalism of Eq. (S6), the replacement (S31) corresponds to replace each $\partial_i^2 \epsilon_{\mathbf{k}}$ term in Eq. (S8) with $\xi_{\mathbf{k}}$, thus removing the divergence of the CP part at $\omega = 2\Delta$ and loosing the tensorial nature of the CP term. This mistake lead the authors of Ref. [S5] to miss the existence of CP processes, and to attribute the THG intensity always to the sub-dominant Higgs-mode excitation.

* lara.benfatto@roma1.infn.it

- [S1] N. Nagaosa, *Quantum Field Theory in Condensed Matter Physics*, Springer, 1999.
- [S2] S. De Palo, C. Castellani, C. Di Castro, and B.K. Chakraverty, Phys. Rev. B 60, 564 (1999).
- [S3] L. Benfatto, A. Toschi, and S. Caprara, Phys. Rev. B. **69**, 184510 (2004) and references therein.
- [S4] T. Cea, C. Castellani, G. Seibold, L. Benfatto, Phys. Rev. Lett. **115**, 157002 (2015).
- [S5] R. Matsunaga, N. Tsuji, H. Fujita, A. Sugioka, K. Makise, Y. Uzawa, H. Terai, Z. Wang, H. Aoki and R. Shimano, Science **345**, 1145 (2014).
- [S6] N. Tsuji and H. Aoki, Phys. Rev. B **92**, 064508 (2015).